

PubMed	Nucleotide	Protein	Genome	Structure	PopSet	Taxonomy	OMIM	Books
Search <u>Nucleotide</u> <input type="text"/> for <input type="text"/>							Go	Clear
Limits		Preview/Index		History		Clipboard		Details
Display	<u>default</u>	<input type="text"/>	Save	Text	Add to Clipboard			

☐ 1: K01760. HSV1 (KOS) glycop...[gi:330082]

Related Sequences, Protein, PubMed, Taxonomy

LOCUS HS1GB 3758 bp DNA linear VRL 02-AUG-1993
 DEFINITION HSV1 (KOS) glycoprotein B gene, complete cds.
 ACCESSION K01760
 VERSION K01760.1 GI:330082
 KEYWORDS glycoprotein; glycoprotein B.
 SOURCE Herpes simplex virus type 1 (strain KOS) DNA, maintained on HEL cells.
 ORGANISM human herpesvirus 1
 Viruses; dsDNA viruses, no RNA stage; Herpesviridae; Alphaherpesvirinae; Simplexvirus.
 REFERENCE 1 (bases 1 to 3758)
 AUTHORS Bzik,D.J., Fox,B.A., DeLuca,N.A. and Person,S.
 TITLE Nucleotide sequence specifying the glycoprotein gene, gB, of Herpes simplex virus type 1
 JOURNAL Virology 133, 301-314 (1984)
 MEDLINE 84174058
 REFERENCE 2 (bases 1 to 3321)
 AUTHORS Bzik,D.J., Debroy,C., Fox,B.A., Pederson,N.E. and Person,S.
 TITLE The nucleotide sequence of the gB glycoprotein gene of HSV-2 and comparison with the corresponding gene of HSV-1
 JOURNAL Virology 155, 322-333 (1986)
 MEDLINE 87071654
 REFERENCE 3 (bases 1 to 3311)
 AUTHORS Pederson,N.
 JOURNAL Unpublished (1987)
 COMMENT [2] revises [1].
 [3] revises [1],[2].
 Draft entry and computer-readable sequence for [3] kindly provided by N.Pederson 14-APR-1987.
 There are 4 antigenically distinct glycoproteins in HSV1 infected cells -- gB, gC, gD and gE. The glycoprotein formerly known as gA is now know to be a precursor to gB and has been renamed pgB. gB is the only glycoprotein that is known to be required for viral growth. The gB mRNA probably runs about from bases 257 to 3300. Strand shown runs in direction opposite to that of the standard HSV map, from mu .370 to .345.
 FEATURES Location/Qualifiers
 source 1..3758
 /organism="human herpesvirus 1"
 /db_xref="taxon:10298"
 CDS 548..3262
 /note="glycoprotein B precursor (pgB)"
 /codon_start=1
 /protein_id="AAA45774.1"
 /db_xref="GI:330083"
 /translation="MHQGAPSWGRRWFVWVALLGLTLGVLVASAAPTSPGTPGVAAATQAANGGPATPAPPPLGAAPTGDPKPKKNKKPKNPTPPRPAGDNATVAAGHATLREHLRDIKAENTDANFYVCPPTGATVVQFEQPRRCPTRPEGQNYTEGIAVVFKENIAPYKFKATMYKDVTVSQVWFGHRYSQFMGIFEDRAPVPFEEVIDKINAKGVCIRSTAKYVRNNLETTAFHRDDHETDMECLKPANAARTSRGWHTTDLKYNPSRVEAFHRYGTTVNCIVEEV"

HSV-1
SB 1

HSV2 SB

DARSVYPYDEFVLATGDFVYMSPFYGYREGSHTTEHTTYAADRFKQVDGFYARDLTTKA
 RATAPTRNLLTTPKFTVAWDWVPKRPSVCTMTKWQEVDEMLRSEYGGSFSSDAIS
 TTFTTNLT EYPLSRVDLGDICGKDARDAMDRI FARRYNATHIKVGQPQYYQANGGFLI
 AYQPLLSNTLAELYVREHLREQSRKPPNPPTPPPPGASANASVERIKTTSSIEFARLQF
 TYNHIQRHVNDMLGRVAIAWCELQNHLELTWNEARKLNPNAIASVTVGRRVSARMLGD
 VMAVSTCVPVAADNVIVQNSMRISSRPGACYSRPLVSFRYEDQGPLVEGQLGENNELR
 LTRDAIEPCTVGHRRYFTFGGGYVYFEEYAYSHQLSRADITTVSTFIDLNITMLEDHE
 FVPLEVYTRHEIKDSGLLDYTEVQRRNQLHDLRFADIDTVIHADANAAMFAGLGAFFE
 GMDGLGRAVGKVMGIVGGVVSASVSGVSSFMSNPF GALAVGLLVLAGLAAFFAFRYV
 MRLQSNPMKALYPLTTKELKNPTNP DASGE GEGGDFDEAKLAEAREMIRYMALVVSAM
 ERTEHKAKKKGTSALLSAKVTDMMRKRNTNYTQVPNKDGDADDDLL"

variation

1364

/note="g in KOS; a in mutant tsJ20"

variation

1676

/note="c in KOS; t in mutant tsJ12"

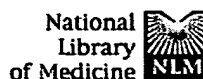
BASE COUNT 717 a 1286 c 1128 g 627 t
 ORIGIN 295 bp upstream of XhoI site; left half of BamHI-G frag at .370 mu.
 1 gtcaacgggc ccctctttga tcaactccacc cacagcttcg cccagccccc caacaccgcg
 61 ctgtattaca gcgtcgagaa cgtggggctc ctgccgcacc tgaaggagga gctcgcccg
 121 ttcacatcatg gggcgggggg ctcgggtgct gattgggccg tcagcgaatt tcagaggttt
 181 tactgttttg acggcatttc cggaataacg cccactcagc gcgccgcctg gcgatatatt
 241 cgcgagctga ttatcgccac cacactcttt gcctcgggtc accggtgcgg ggagctcgag
 301 ttgcgcggcc cggactgcag ccgcccgcacc tccgaaggtc gttaccgta cccgcccggc
 361 gtatatctca cgtacgactc cgactgtccg ctggtggcca tcgtcgagag cgccccgcac
 421 ggctgtatcg gcccccggtc ggtcgtggtc tacgaccgcg acgttttctc gatcctctac
 481 tcggtcctcc agcacctcgc ccccaggcta cctgacgggg ggacgacgg gccccgtag
 541 tcccgcgatg caccaggggc cccctcgtg ggggcgcgg tggttcgtcg tatgggcgt
 601 cttgggggtg acgctggggg tcctggtggc gtcggcggt cgcacttccc cggcacgccc
 661 tggggctcgc gcgcgaccc aggcgcgcaa cggggggcct gccactccg gcgcgcggc
 721 ccttggcgccc gcccacacgg gggacccgaa accgaagaag aacaaaaaac cgaaaaaccc
 781 aacgccacca cgccccgcg gcgacaacgc gaccgtcgcc gcgggccacg ccaccctgcg
 841 cgagcacctg cgggacatca aggcggagaa caccgatgca aacttttacg tgtgccacc
 901 cccacggggc gccacggtgg tgcagttcga gcagccgcg cgctgcccga cccggcccga
 961 gggtcagaac tacacggagg gcatcgcggt ggtcttcaag gagaacatcg ccccgtaaa
 1021 gttcaaggcc accatgtact acaaagacgt caccgtttcg cagggtgtgt tcggccaccg
 1081 ctactcccag tttatgggga tctttgagga ccgcgcccc gtccccctcg aggaggtgat
 1141 cgacaagatc aacgccaagg gggctctgctg gtccacggcc aagtacgtg gcaacaacct
 1201 ggagaccacc gcgtttcacc gggacgacca cgagaccgac atggagctga aaccggccaa
 1261 cgccgcgacc cgacgagacc ggggctggca caccaccgac ctcaagtaca acccctcgcg
 1321 ggtggaggcg ttccaccggt acgggacgac ggtaaaactgc atcgtcgagg aggtggacgc
 1381 gcgctcggtg taccgtacg acgagtttgt gctggcgact ggcgactttg tgtacatgtc
 1441 cccgttttac ggctaccggg aggggtcgca caccgaacac accacgtacg ccgccgaccg
 1501 cttcaagcag gtgcgacggt tctacgcgcg cgacctcacc accaaggccc gggccacggc
 1561 gccgaccacc cggaacctgc tcacgacccc caagttcacc gtggcctggg actgggtgccc
 1621 aaagcggccc tcggtctgca ccatgaccaa gtggcaggaa ttggacgaga tgcgtcgctc
 1681 caggtacggc ggctccttcc gattctcctc cgacgccata tccaccacct tcaccaccaa
 1741 cctgaccgag taccgctctc cgcgctgga cctgggggac tgcacggca aggacgcccg
 1801 cgacgcatat gaccgcatct tcgcccgcag gtacaacgcg acgcacatca aggtgggcca
 1861 gccgcagtac taccaggcca atgggggctt tctgatcgcg taccagcccc ttctcagcaa
 1921 cacgctcgcg gagctgtacg tgcgggaaca cctccgagag cagagccgca agccccaaa
 1981 cccacgccc ccgcccggcg gggccagcgc caacgcgtcc gtggagcgca tcaagaccac
 2041 ctccctccatc gagttcgccc ggctgcagtt tacgtacaac cacatacagc gccatgtcaa
 2101 cgatattgtt ggcgcggtt ccatcgctg gtgcgagcta cagaatcacg agctgaccct
 2161 gtggaacgag gcccgcaagc tgaaccccaa gcctatcgcc tcgggtcaccg ttggcggcg
 2221 ggtgagcgcg cggatgctcg gcgacgtgat ggccgtctcc acgtgcgtgc cgtcgccgc
 2281 ggacaacgtg atcgtccaaa actcgtatgc catcagctcg cgccccggg cctgctacag
 2341 ccgccccctg gtcagctttc ggtacgaaga ccaggggccc ttggtcgagg ggcagctggg
 2401 ggagaacaac gagctgcggc tgacgcgcga tgcgatcgag ccgtgcaccg tgggacaccg
 2461 gcgctacttc accttcgggtg ggggctacgt gtacttcgag gagtacgcgt actcccacca
 2521 gctgagccgc gccgacatca ccaccgtcag caccttcacg gacctcaaca tcaccatgct
 2581 ggaggatcac gagtttgtcc ccctggagg gtacacccgc cacgagatca aggacagcg
 2641 cctgctggac tacacggagc tccagcgccg caaccagctg cacgacctgc gcttcgcga
 2701 catcgacacg gtcacccacg ccgacgccaa cgccgccatg ttcgcgggcc ttggcgcgctt
 2761 cttcgagggg atgggcgacc tggggcgcg ggtcggaag gtggtgatgg gcatcgtggg
 2821 cggcgtggtg tcggccgtgt cgggcggtc ctccttcacg tccaacccct ttggggcgct

```
2881 ggccgtgggt ctgttggtcc tggccggcct ggcggcggcc ttcttcgcct ttcgttacgt
2941 catgcggctg cagagcaacc ccatgaaggc cctgtaccct ctaaccacca aggagctcaa
3001 gaaccccacc aacccgacg cgtccgggga gggcgaggag ggcggcgact ttgacgaggc
3061 caagctagcc gaggccaggg agatgatacg gtacatggcc ctggtgtcgg ccatggagcg
3121 cacggaacac aaggccaaga agaagggcac ggcgcgctg ctcagcgcca aggtcaccga
3181 catggtcatg cgcaagcgcc gcaacaccaa ctacaccaa gttcccaaca aagacggtga
3241 cgccgacgag gacgacctgt gacggggggg ttgttgtaaa taaaaaccac ggggtgttaa
3301 ccgcatgcgc atcttttggt ttttttggtt ggtcagcctt ttgtgtgtgt gtgggaagaa
3361 agaaaaagga acacataaac tccccgggt gtccgcggcc tgtttcctct ttcctttccc
3421 gtgacaaaac ggaccccctt ggtcagtgcc gatttcctcc cccccacgcc ttcctccacg
3481 tcaaaggctt ttgcattgta aagctaccgg cctaccgcc cctcccaata aaaaaaaaag
3541 aacatacacc aatgggtctt atttggtatt acctggttta tttaaaaaga tatacagtaa
3601 gacatcccat ggtaccaaag accggggcga atcagcgggc ccccatcatc tgagagacga
3661 acaaatcggc ggcgcgggcc gtgtcaacgt ccacgtgtgc tgcgctgctg gcgttgacaa
3721 ggccccggcc tccgcgttgg atgctccggt tgggatcc
```

//

Revised: October 24, 2001.

[Disclaimer](#) | [Write to the Help Desk](#)
[NCBI](#) | [NLM](#) | [NIH](#)



PubMed	Nucleotide	Protein	Genome	Structure	PopSet	Taxonomy	OMIM	Books
Search PubMed	[v] for						Go	Clear
Limits		Preview/Index		History		Clipboard		Details

Entrez
PubMed

☐ 1: Virology 1984 Mar;133(2):301-14

Related Articles, Nucleotide, Protein, ^{NEW} Book:
LinkOut

PubMed
Services

Nucleotide sequence specifying the glycoprotein gene, gB, of herpes simplex virus type 1.

Bzik DJ, Fox BA, DeLuca NA, Person S.

Related
Resources

The nucleotide sequence thought to specify the glycoprotein gene, gB, of the KOS strain of herpes simplex virus type 1 (HSV-1) has been determined. A 3.1-kilobase (kb), viral-specified RNA was mapped to the left half of the BamHI-G fragment (0.345 to 0.399 map units). TATA, CAT-box, and possible mRNA start sequences characteristic of HSV-1 genes are found near 0.368 map units. The first available ATG codon is at 0.366 and the first in-phase chain terminator at 0.348 map units. A polyA-addition signal (AATAAA) occurs 17 nucleotides past the chain terminator. Translation of these sequences would yield a 100.3-kilodalton (kDa) polypeptide characterized by a 5' signal sequence, nine N-linked saccharide addition sites, a strongly hydrophobic membrane-spanning sequence, and a highly charged 3' cytoplasmic anchor sequence. Two mutants of KOS, tsJ12 and tsJ20, that are temperature-sensitive for viral growth and for the production of gB, have been physically mapped to 0.357 to 0.360 and 0.360 to 0.364 map units, respectively (DeLuca et al., in preparation). The nucleotide sequence of the mutants was determined in these regions. In both cases a single amino acid replacement within the 100.3-kDa polypeptide is predicted from the sequence analysis.

MeSH Terms:

- Amino Acid Sequence
- Base Sequence
- Cell Line
- DNA Restriction Enzymes
- DNA, Recombinant/analysis
- Genes, Structural*
- Genes, Viral*
- Human
- Lung/embryology
- Nucleic Acid Hybridization
- Simplexvirus/genetics*
- Support, Non-U.S. Gov't
- Support, U.S. Gov't, P.H.S.

Substances:

- DNA Restriction Enzymes
- DNA, Recombinant

Secondary source id:


- GENBANK/K01760

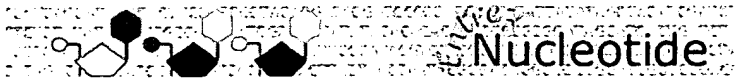
PMID: 6324454 [PubMed - indexed for MEDLINE]



[Write to the Help Desk](#)
[NCBI](#) | [NLM](#) | [NIH](#)
[Department of Health & Human Services](#)
[Freedom of Information Act](#) | [Disclaimer](#)

sparc-sun-solaris2.8 Feb 7 2002 10:32





PubMed

Nucleotide

Protein

Genome

Structure

PopSet

Taxonomy

OMIM

Books

Search

Nucleotide

▼

for

Go

Clear

Limits

Preview/Index

History

Clipboard

Details

Display

default

▼

Save

Text

Add to Clipboard

☐ 1: M21629. Herpes simplex vi...[gi:330109]

Related Sequences, Protein, PubMed, Taxonomy

LOCUS HS1GLYB 9756 bp DNA linear VRL 01-JUL-1994

DEFINITION Herpes simplex virus type 1 glycoprotein B gene (gB-1), complete

cds.

ACCESSION M21629

VERSION M21629.1 GI:330109

KEYWORDS glycoprotein B.

SOURCE Herpes simplex virus type 1 DNA.

ORGANISM human herpesvirus 1

Viruses; dsDNA viruses, no RNA stage; Herpesviridae;

Alphaherpesvirinae; Simplexvirus.

REFERENCE 1 (bases 1 to 9756)

AUTHORS Hammerschmidt,W., Conraths,F., Mankertz,J., Pauli,G., Ludwig,H. and

Buhk,H.J.

TITLE Conservation of a gene cluster including glycoprotein B in bovine

herpesvirus type 2 (BHV-2) and herpes simplex virus type 1 (HSV-1)

JOURNAL Virology 165 (2), 388-405 (1988)

MEDLINE 88306231

FEATURES

source Location/Qualifiers

1..9756

/organism="human herpesvirus 1"

/db_xref="taxon:10298"

gene 6555..9266

/gene="gB-1"

CDS 6555..9266

/gene="gB-1"

/codon_start=1

/product="glycoprotein B"

/protein_id="AAA19496.1"

/db_xref="GI:330110"

/translation="MRQGAARGCRWFVWALLGLTLGVLVASAAPSSPGTPGVAAATQ

AANGGPATPAPPAPGPAPTGDTPKPKNKKPKNPPPPRPAGDNATVAAGHATLREHLRD

IKAENTDANFYVCPPTGATVVQFEQPRRCPTRPEGQNYTEGIAVVFKENIAPYKFKA

TMYKDVTVSQVWFGHRYSQFMGIFEDRAPVPFEEVIDKINAKGVCRSTAKYVRNNLE

TTAFHRDDHETDMELKPANAATRTSRGWHTTDLKYNPSRVEAFHRYGTTVNCIVEEVD

ARSVYPYDEFVLATGDFVYMSPFYGYREGSHTEHTSYAADRFKQVDGIFYARDLTTKAR

ATAPTTRNLLTTPKFTVAWDWVPKRPSVCTMTKWQEVDEMLRSEYGGSFRESSDAIST

TFTTNLTLEYPLSRVDLGDGIGKDARDAMDRIFARRYNATHIKVGQPQYYLANGGFLIA

YQPLLSNTLAELYVREHLREQSRKPPNPPTPPPPGASANASVERIKTTSSIEFARLQFT

YNHIQRHVNDMLGRVAIAWCELQNHETLWNEARKLNPNAIASATVGRRVSARMLGDV

MAVSTCVPAADNVIVQNSMRISSRPGACYSRPLVSFRYEDQGPLVEGQVGENNELRL

TRDAIEPCTVGHRRYFTFGGGYVYFEEYAYSHQLSRADITTVSTFIDLNITMLDHEF

VPLEVYTRHEIKDSGLLDYTEVQRRNQLHDLRFADIDTVIHADANAAMFAGLGAFEG

MGDLGRAVGKVVMGIVGGVSAVSGVSSFMSNPFALAVGLLVLAGLAAFFAFRYVM

RLQSNPMKALYPLTTKELKNPTNPDASGEGEGGDFDEAKLAEAREMIRYMALVSAME

RTEHKAKKKGTSALLSAKVTDMVMRKRNTNYTQVPNKDGADEDDL"

BASE COUNT 1662 a 3321 c 3202 g 1571 t

ORIGIN

1 agtgtacata taccaaccgc ataccagacg caccgcaccc ggagcacctg accgtaagca

61 tctgtgcctc tcgcagggac cccgcgttgc cggccgccgg ggttcatcgg cgccccctgt

121 gttaccgccg gggttgctcg tgaaggggag ggattcattc cccaaccgcc gtctccaacc

```
181 ctcccccttga ccgtcgcccgc ccccccccg gatttgacgc tcgggagaca taccttgctc
241 ggcgtccgctg gtcgtgccgg gattacctcc gttcgcggac cgattgacaa aaggacatgg
301 agacaaagcc caagacggca accaccatca aggtccccc cgggcccctg ggatacgtgt
361 acgctcgcgc gtgtccgtcc gaaggcatcg agcttctggc gttactgtcg gcacgcagcg
421 gcgattccga cgtcgccgtg gcgcccctgg tcgtgggcct gaccgtggag agcggctttg
481 aggccaaact ggccgtggtc gtgggttctc gcacgacggg gtcgggggt accgcggtgt
541 ccctgaaact gacgccctcg cactacagct cgtccgtgta cgtctttcac ggcgcccgcc
601 acctggaccc cagcacccag gccccgaacc tgacgcgcct ttgcgagcgg gcacgccgcc
661 attttggttt ttccgactac acccccgggc ccggcgacct caaacacgag acgacggggg
721 aggcgtgtgt tgagcgctc ggctggacc cggaccgcgc cctcctgtat ctggtcggtt
781 ccgagggttt caaggaggcc gtgtcgatca acaacacctt tctgcacctg ggaggctcgg
841 acaaggtaac cataggcggg gcggaggtgc accgcatacc cgtgtaccgg ttgcagctgt
901 tcatgccgga ttttagccgt gtcatcgtag agccgttcaa cgccaaccac cgatcgatcg
961 gggagaaaatt tacctaccgg cttccgtttt ttaaccgccc cctcaaccgc ctctgttcg
1021 aggcgtcgtt gggaccgcc gccgtggcac tgcgatgccg aaacgtggac gccgtggccc
1081 gcgcggccgc ccacctggcg tttgacgaaa accacgaggg cgccgccctc cccgccgaca
1141 ttacgttcac ggccttcgaa gccagccagg gtaagacccc gcggggcgag cgcgacggcg
1201 gcggcaaggg cgcggcgggc gggttcgaac agcgctggc ctccgtcatg gccggagacg
1261 ccgcctggc cctcgagtct atcgtgtcga tggcgtctt tgacgagccg cccaccgaca
1321 tctccgctg gccgctgttc gagggccagg acacggccgc ggcccgcgc ggcccgctcg
1381 gggcctacct ggcgcgcgcc gcgggactcg tggggccat ggtatttagc accaactcgg
1441 ccctccatct caccgaggtg gacgacgccg gccggcgga cccaaaggac cacagcaaac
1501 cctcctttta ccgcttcttc ctctgcccc ggaccacgt ggcgccaac ccacaggtgg
1561 accgcgaggg acacgtggtg cccgggttcg agggtcggcc caccgcgcc ctctcgcgcg
1621 gaaccaagga atttgccggc gagcacctgg ccattgctgt tgggttttcc ccggcgctgc
1681 tggccaagat gctgttttac ctggagcgct gcgacggcgc cgtgatcgtc gggcgccagg
1741 agatggacgt gtttcgatac gtcgcggact ccaaccagac cgacgtgcc tgaacctat
1801 gcaccttcga cacgcgccac gccctgcgat acacgacgt catgcgcctc cgggcgcgcc
1861 atccaaagt cccagcgccc ccccgcgagg ccacggcgt ctcgggacc atgaacagca
1921 tgtacagcga ctgcgacgtg ctgggaaact acgcccctt ctcggccctg aagcgcgcgg
1981 acgatccga gaccgcccg accatcatgc aggagacgta ccgcgcggcg accgagcgcg
2041 tcatggccga actcgagacc ctgcagtac tggaccaggc ggtccccac gccatggggc
2101 ggctggagac catcatcacc aaccgcgagg ccctgcatac ggtggtgaac aacgtcaggc
2161 aggtcgtgga ccgcgaggtg gagcagctga tgcgcaacct ggtggagggg aggaacttca
2221 agtttcgca cggtctgggc gaggccaacc acgccatgtc cctgacgtg gaccgtacg
2281 cgtgcgggcc gtgccccctg cttcagcttc tcgggcggcg atccaacctc gccgtgtacc
2341 aggacctgaa cctgagtcag tgccagggg gtttcgcccg gcagtcggtc cgcgtgcga
2401 actttcgcaa tcaattccaa ccggtgctgc ggcggcgct gatggacatg tttaaacacg
2461 ggtttctgtc ggccaaaacg ctgacggtcg cgctctcgga gggggcggt atctgcgcc
2521 ccagcctaac ggccggccag acggcccccg ccgagagcag cttcgagggc gacgttgccc
2581 gcgtgaccct ggggtttccc aaggagctgc gcgtcaagag ccgcgtgttg ttcgcggcg
2641 cgagcgccaa cgcgtccgag gccgccaagg cgcgggtcgc cagcctccag agcgcctacc
2701 agaagcccga caagcgcgtg gacatcctcc tcggaccgct gggctttctg ctgaagcagt
2761 tccacgcggc catcttcccc aacgcaaac ccccggggtc caaccagccg aaccgcagct
2821 ggttctggac ggccctccaa cgcaaccagg ttcccgcggc gctcctgtcg cgcgagaca
2881 tcgagaccat cgcgttcatt aaaaagtttt ccctggacta cggcgcgata aactttatta
2941 acctggcccc caacaacgtg agcgagctgg cgatgtacta catggcaaac cagattctgc
3001 ggtactgcga tcaactcgaca tacttcatca acacccttac ggccatcatc gcgggtccc
3061 gccgtcccc cagcgtgcag gctgcggccg cgtggtccgc gcaggcggg gcgggcttg
3121 aggcggggg ccgcgcgctg gtggacgccg tggacgcgca tccgggcgcg tggacgtcca
3181 tgttcgccag ctgcaacctg ctgcggcccc tcattggcgg cgcgcccatg gtcgtgttg
3241 ggttgagcat cagcaagtac tacggcatgg ccggcaacga ccgtgtgttt caggccggga
3301 actgggccag cctgatgggc ggcaaaaacg cgtgcccgct ccttattttt agccgacccc
3361 gcaagtctgt cctggcctgt ccccgggcgg gtttgtgtg cgcgccctcg gacttcggc
3421 cgggagcgcg cgaaagctcg ctgtgcgagc agtccgggg cattatctcc gagggcgggg
3481 cgccgctcgc cagtagcgtg ttcgtggcga ccgtgaaaag cctggggccc cgcacccaac
3541 agctgcagat cgaggactgg ctggcgctcc tggaggacga gtacctaacg gaggagatga
3601 tggagctgac cgcacgtgcc ctggagcgcg gcaacggcga gtggtcgacg gacgcggccc
3661 tggaggtggc gcacgaggcc gaggccctag tcagccaact cggcaacgcc ggggaggtgt
3721 ttaactttgg ggattttggc tgcgaggacg acaacgcgac gccgttcggc ggccccgggg
3781 ccccgggacc ggcatttgcc ggccgcaaac gggcgttcca cgggatgac ccgtttgggg
3841 aggggcccc cgacaaaaag ggagactga cgttgatat gctgtgagg gttggggg
3901 gggggaacct agggcggggc ggggaatgt gttaaaataa attattgcta cgacatccgt
3961 gcttggttgt gttccgtgtc tatactctct ggcgggcgtt gattcctctc cgcggtgtct
4021 gggaatagaa cagaaacgca cgcgcgcgcg actcccggct tgccggtcgg cgggcccgcg
```

```
4081 ggaggccgcc ccgaataggg ggaccccggg gctcagccag acgccgggta gcgtacggac
4141 cgctcggctg caccgcatac tccggccggc gtacagatcg gcgccgcgag atggccgccc
4201 cgtgtgccga gccaccgtg ccccgtaaaa agttgttagc cctgctcgga cagggtcgaga
4261 cctatgtgtt tcagatagag ctgctccggc ggtgcgaccc ccacatcgga cgggggaagc
4321 tcccccaact gaagctgaac gcgcttcagg tgcgggcgct gcggcgctcg ctgaggccgg
4381 gcctggaggc ccaggccggg gcctttctca ccccgctgtc ggtcaccctg gagtgtctgc
4441 tagagtacgc gtggcgcgag ggcgagcggc tcctgggcag cctggagacg ttcgcgaccg
4501 cgggagacgt cgcggtgttt ttcacggaga ccatgggcct ggcccgaccc tgtccgtatc
4561 accaacgggt caggctggat acgtatggcg ggaccgtcca tatggagctg tgtttcctgc
4621 acgacgtcga gaactttcta aagcagctaa actactgcca cctcatcacc ccctcgcgcg
4681 gcgccaccgc cgcgttggag cgcgttcggg agtttatggt gggggcggtg gggtcgggcc
4741 ttatcgtccc cccggagctt agcgaccgtg cccaccctg cgcggtctgt ttcgaggaac
4801 tgtgtgtgac ggcgaaccag gggcgacga tgcggcgccg cctggcgga cgtatctgta
4861 accacgtcac ccagcaggcg cagggtcggc tggacgcca cgagctcgcg cggtagctgc
4921 cccacgccgc cgggctgtcg gacgccgacc gcgcgcgggc gctctccgtg ttggaccatg
4981 cgctggcccg gaccgcgggg ggcgacgggc agccccaccc gtcggccgag aacgactcgg
5041 tccgcaagga ggccgacgcc ctgctggagg cgcacgacgt gtttcaggcc accacgcccg
5101 gcctgtacgc catcagcgaa ttgcgattct ggctcgcgtc cggcgaccgc gccggccaga
5161 ccaccatgga cgcgtttgcc agcaacctga ccgcgtggc gcggcgcgag ttgcagcagg
5221 agaccgccgc ggtggccgtg gaactggcgc tgttcgggag cggggcgga catttcgatc
5281 gcgcgttcgg gagccacctg gcggcgctgg atatggtgga cggcctaata atcggcggtc
5341 aggccacgtc acccgacgat cagatcgagg cgctcatccg cgctgtctac gaccaccacc
5401 tgacgacgcc gctcttgccg gcctcgtca gccccgaaca gtgcgacgag gaggcgctgc
5461 gtcgcgtgct ggcgcgcatg ggggcggggg gcgcggcgga cgcgccaaag ggcgcgcgcg
5521 gcccgacga cgacggggac cgtgtcgccg tagaggaagg ggcacggggg ttgggagctc
5581 ccgggggagg gggcgaggac gaagaccgtc gcccggggcc cggggggcag gggcccgaga
5641 cgtgggggga catcgccacg caagcgcccg cggacgtgcg ggagcgacgg cggctgtacg
5701 cggaccgcct gacgaagcgg tcgttgccca ccctggggcg ctgctccgc gagcagcgcg
5761 cggagctcga gaagatgctg cgggtcagcg tccacggcga gttgctgcc gcagcgttcg
5821 ccgcggtcgc caacggcttt gcggcgcgcg cgcgcttctg cgccctgacg gcgggcgcg
5881 gcacggtcat cgacaaccgc tcggcgcccg gcgtgttcga cgcgcaccgg ttcagcgag
5941 cgtctctcct gcgacaccag gtggaccgg cctgtctccc cagcatcacc catcgcttct
6001 tcgagctcgt caacgggccc ctctttgatc actccacca cagcttcgcc cagcccccca
6061 acaccgcgct gtattacagc gtcgagaacg tggggctcct gccgcacctg aaggaggagc
6121 tcgcccgggt catcatgggg gcgggggctc ggggtctgat tgggccgtca gcgaatttca
6181 gaggttttat tgttttgacg gcatttccgg aataacgccc actcagcgcg ccgctggcg
6241 atatgttcgc gagctgatta tcggccacc acctcttgcc tcggtctacc ggtcgggga
6301 gctcgagtgt cggcgcccgg actgcagccg cccgacctcc gaaggctcgtt accgttacct
6361 gcccgcgcta tatctcacgt acgactccga ctgtccgctg gtggccatcg tcgagagcgc
6421 ccccgacggc tgtatcggcc cccggtcggg cgtggtctac gaccgagacg ttttctcgat
6481 cctctactcg gtccctccagc acctcgcccc caggctacct gacggggggc acgacggggc
6541 cccgtagtcc cgccatgcgc cagggcgccg cgcgggggtg ccggtggttc gtcgtatggg
6601 cgctcttggg gttgacgctg ggggtcctgg tggcgtcggc ggctccgagt tccccggca
6661 cgcttggggg cgcggccgcg acccaggcga cgaacggggg acctgccact ccggcgccgc
6721 ccgcccctgg ccccgcccca acgggggaga cgaaaaccga gaagaacaaa aaaccgaaaa
6781 accaacggcc cccgcgcccc gccgcggaca acgcgaccgt cgcccggggc cagccacccc
6841 tgcgcgagca cctgcgggac atcaaggcgg agaacaccga tgcaaacttt tacgtgtgcc
6901 cccccccac gggcgccacg gtggtgcagt tcgagcagcc gcgcgctgc ccgaccggc
6961 ccgagggtca gaactacacg gagggcatcg cggtggtctt caaggagaac atcgccccgt
7021 acaagttcaa ggccaccatg tactacaaag acgtcaccgt ttcgaggtg tggttcggcc
7081 accgctactc ccagtttatg gggatctttg aggaccgcgc ccccgctccc ttcgaggagg
7141 tgatcgacaa gatcaacgcc aagggggtct gtcggtccac ggccaagtac gtgcgcaaca
7201 acctggagac caccgcgttt caccgggacg accacgagac cgacatggag ctgaaccgg
7261 ccaacggcgc gaccgcacg agccggggtc ggacaccac cgacctcaag tacaaccctt
7321 cgcgggtgga ggcgttcac cggtagggga cgacggtaaa ctgcatcgtc gaggagggtg
7381 acgcgcgctc ggtgtacccg tacgacgagt ttgtgctggc gactggcgac tttgtgtaca
7441 tgtccccgtt ttacggctac cgggaggggt cgcacaccga acacaccagc tacgccggcg
7501 accgcttcaa gcaggttgac ggcttctacg cgcgcgacct caccaccaag gccggggcca
7561 cggcgccgac cccccggaac ctgctcacga cccccaagtt caccgtggcc tgggactggg
7621 tgccaaagcg cccgtcggtc tgcaccatga ccaagtggca ggaggtggac gagatgctgc
7681 gctccgagta cggcggtccc ttccgattct cctccgacgc catatccacc accttaccac
7741 ccaacctgac cgagtacccg ctctcgcgct tggacctggg ggactgcacg ggcaaggacg
7801 cccgcgacgc catggaccgc atcttcgccc gcaggtacaa cgcgacgcac atcaagggtg
7861 gccagccgca gtactacctg gccaatgggg gctttctgat cgcgtaccag ccccttctca
7921 gcaacacgct cgcgagctg tacgtgcggg aacacctccg agagcagagc cgcaagcccc
```



```
7981 caaaccccccac gcccccgccg cccggggcca gcgccaacgc gtccgtggag cgcacaaaga
8041 ccacctcctc catcgagttc gcccggtgc agtttacgta caaccacata cagcgccatg
8101 tcaacgatat gttgggcccgc gttgccatcg cgtggtgcga gctgcagaat cacgagctga
8161 ccctgtggaa cgaggccccgc aagctgaacc ccaacgccat cgcctcgcc accgtgggccc
8221 ggcgggtgag cgcgcggatg ctcggcgacg tgatggccgt ctccacgtgc gtgccggctg
8281 ccgcggacaa cgtgatcgtc caaaactcga tgcgcatcag ctcgcgcccc ggggcctgct
8341 acagccgccc cctggtcagc ttctcggtacg aagaccaggg cccgttggtc gaggggcagg
8401 tgggggagaa caacgagctg cggctgacgc gcgatgcgat cgagccgtgc accgtgggac
8461 accggcgcta cttcaccttc ggtgggggct acgtgtactt cgaggagtac gcgtactccc
8521 accagctgag ccgcgccgac atcaccaccg tcagcacctt catcgacctc aacatcacca
8581 tgctggagga tcacgagttt gtccccctgg aggtgtacac ccgccacgag atcaaggaca
8641 gcggcctgct ggactacacg gaggtccagc gccgcaacca gctgcacgac ctgccttcg
8701 ccgacatcga cacggtcac cagccgacg ccaacgccgc catgttcgag ggcctgggag
8761 cgttcttcga ggggatgggc gacctggggc gcgcggtcgg caaggtggtg atgggcatcg
8821 tgggcggcgt ggtatcgccc gtgtcgggag tgctctcctt catgtccaac ccctttgggg
8881 cgctggccgt gggctctgtt gtcttgccg gcctggcggg ggctttcttc gcctttcgt
8941 acgtcatgag gctgcagagc aaccccatga aggcctgta cccgctaacc accaaggagc
9001 tcaagaaccc caccaacccg gacgcgtccg gggagggcga ggagggcggc gactttgacg
9061 aggccaagct agccgaggcc cgggagatga tacggtacat ggccctggtg tctgccatgg
9121 agcgcacgga acacaaggcc aagaagaagg gcacgagcgc gctgctcagc gccaaagtca
9181 ccgacatggt catgcgcaag cgccgcaaca ccaactacac ccaagttccc aacaaagacg
9241 gtgacgccga cgaggacgac ctgtgacggg gggtttgttg taaataaaaa ccacgggtgt
9301 taaaccgcat gtgcatcttt tggtttgtt gtttggtagc ctttttgtgt gtgtgggaag
9361 aaagaaaagg gaacacataa actcccccg gtgtccgagg cctgtttcct ctttcctttc
9421 ccgtgacaaa acggaccccc ttggtcagtg ccgattcccc cccacgcct tcctccacgt
9481 cgaaggcttt tgcattgtaa agctaccgcg ctaccgcgc ctcccaataa aaaaagaaca
9541 tacaccaatg ggtcttattt ggtattacct ggtttattta aaaagatata cagtaagaca
9601 tcccatggta ccaaagaccg gggcgaatca gcgggcccc atcatctgag agacgaacaa
9661 atcggcggcg cgggccgtgt caacgtccac gtgtgctgag ctgctggcgt tgacaagggc
9721 cccggcctcc gcgttgatg cctccggtt ggatcc
```

//

Revised: October 24, 2001.

[Disclaimer](#) | [Write to the Help Desk](#)
[NCBI](#) | [NLN](#) | [NIH](#)